

# The standardisation process failed during the COVID-19 exams fiasco, but so too did the policy process



*In the summer of 2020, after cancelling exams, the UK and devolved governments sought teacher estimates on students' grades, but supported an algorithm to standardise the results. When the results produced a public outcry over unfair consequences, they initially defended their decision but reverted quickly to teacher assessment. These experiences, argue [Sean Kippin](#) and [Paul Cairney](#), highlight the confluence of events and choices in which an imperfect and rejected policy solution became a 'lifeline' for four beleaguered governments.*

In 2020, the UK and devolved governments performed a 'U-turn' on their COVID-19 school exams replacement policies. The experience was embarrassing for education ministers and damaging to students. There are significant differences between (and often within) the four nations in terms of the structure, timing, weight, and relationship between the different examinations. However, in general, the A-level (England, Northern Ireland, Wales) and Higher/ Advanced Higher (Scotland) examinations have similar policy implications, dictating entry to further and higher education, and influencing employment opportunities. The [Priestley review](#), commissioned by the Scottish Government after their U-turn, described this as an 'impossible task'.

Initially, each government defined the new policy problem in relation to the need to 'credibly' replicate the purpose of exams to allow students to progress to tertiary education or employment. All four quickly announced their intentions to allocate in some form grades to students, rather than replace the assessments with, for example, remote examinations. However, mindful of the long-term credibility of the examinations system and of ensuring fairness, each government opted to maintain the qualifications and seek a similar distribution of grades to previous years. A key consideration was that UK universities accept large numbers of students from across the UK.

One potential solution open to policymakers was to rely solely on teacher grading (CAG). [CAGs are](#) 'based on a range of evidence including mock exams, non-exam assessment, homework assignments and any other record of student performance over the course of study'. Potential problems included the risk of high variation and discrepancies between different centres, the potential overload of the higher education system, and the tendency for teacher predicted grades to reward already privileged students and punish disabled, non-white, and economically deprived children.

A second option was to take CAGs as a starting point, then use an algorithm to produce 'standardisation', which was potentially attractive to each government as it allowed students to complete secondary education and to progress to the next level in similar ways to previous (and future) cohorts. Further, an emphasis on the technical nature of this standardisation, with qualifications agencies taking the lead in designing the process by which grades would be allocated, and opting not share the details of its algorithm were a key part of its (temporary) viability. Each government then made similar claims when defending the problem and selecting the solution. Yet this approach reduced both the debate on the unequal impact of this process on students, and the chance for other experts to examine if the algorithm would produce the desired effect. Policymakers in all four governments assured students that the grading would be accurate and fair, with teacher discretion [playing a large role](#) in the calculation of grades.

To these governments, it appeared at first that they had found a fair and efficient (or at least defensible) way to allocate grades, and public opinion did not respond negatively to its announcement. However, these appearances proved to be profoundly deceptive and vanished on each day of each exam result. The Scottish national mood shifted so intensely that, after a few days, pursuing standardisation no longer seemed politically feasible. The intense criticism centred on the unequal level of reductions of grades after standardisation, rather than the unequal overall rise in grade performance after teacher assessment and standardisation (which advantaged poorer students).

Despite some recognition that similar problems were afoot elsewhere, this shift of problem definition did not happen in the rest of the UK until (a) their published exam results highlighted similar problems regarding the role of previous school performance on standardised results, and (b) the Scottish Government had already changed course. Upon the release of grades outside Scotland, it became clear that downgrades were also concentrated in more deprived areas. For instance, in [Wales, 42% of students](#) saw their A-Level results lowered from their Centre Assessed Grades, with the figure close to a third for Northern Ireland.

Each government thus faced similar choices between defending the original system by challenging the emerging consensus around its apparent unfairness; modifying the system by changing the appeal system; or abandoning it altogether and reverting to solely teacher assessed grades. Ultimately, all three governments followed the same path. Initially, they opted to defend their original policy choice. However, by 17 August, the UK, Welsh, and Northern education secretaries announced (separately) that examination grades would be based solely on CAGs – unless the standardisation process had generated a higher grade (students would receive whichever was highest).

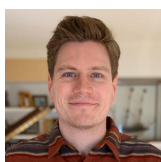
Scotland's initial experience was instructive to the rest of the UK and its example provided the UK government with a blueprint to follow (eventually). It began with a new policy choice – reverting to teacher assessed grades – sold as fairer to victims of the standardisation process. Once this precedent had been set, a different course for policymakers at the UK level became difficult to resist, particularly when faced with a similar backlash. The UK's government's decision in turn influenced the Welsh and Northern Irish governments.

In short, we can see that the particular ordering of choices created a cascading effect across the four governments which created initially one policy solution, before triggering a U-turn. This focus on order and timing should not be lost during the inevitable inquiries and reports on the examinations systems. The take-home message is to not ignore the policy process when evaluating the long-term effect of these policies. Focus on why the standardisation processes went wrong is welcome, but we should also focus on why the policymaking process malfunctioned, to produce a wildly inconsistent approach to the same policy choice in such a short space of time. Examining both aspects of this fiasco will be crucial to the grading process in 2021, given that governments will be seeking an alternative to exams for a second year.

---

Note: the above draws on the authors' [published work](#) in *British Politics*.

### About the Authors



**Sean Kippin** is Lecturer at the University of Stirling.



**Paul Cairney** is Professor at the University of Stirling.

Photo by [Chris Liverani](#) on [Unsplash](#).